U.S. Department of Health and Human Services (HHS)
National Institutes of Health (NIH)
Office of the Director (OD)
Division of Program Coordination, Planning, and Strategic Initiatives (DPCPSI)
Office of Data Science Strategy (ODSS)

Toward an Ethical Framework for Artificial Intelligence in Biomedical and Behavioral Research: Transparency for Data and Model Reuse Workshop

January 31 - February 2, 2024
Rockledge II, 6701 Rockledge Drive
Bethesda, MD
& Virtual

Executive Summary

Federally funded open science often involves the reuse of data and models. NIH researchers who develop and use artificial intelligence (AI) want to advance its responsible use, but they lack guidance for supporting downstream responsible and ethical reuse. NIH's data science and AI initiatives aim to address the ethical, legal, and social implications of the use of AI in healthcare. This three-day hybrid workshop was sponsored by the NIH Office of Data Science Strategy (ODSS), which leads the implementation of the NIH Strategic Plan for Data Science through scientific, technical, and operational collaboration with the institutes, centers, and offices that comprise NIH. The workshop was convened to explore and assess the landscape of responsible AI[1] in the biomedical and behavioral research context and take steps toward improving practice. Experts gathered and shared input on opportunities and challenges with respect to the responsible use and reuse of data and models in the AI lifecycle.

In this workshop, transparency served as an entry point into responsible AI. This workshop took a practical and stakeholder-based approach to transparency, defining transparency as *providing the information that stakeholders need to make informed, responsible, and ethical decisions for data and model reuse*. Participants identified stakeholders across the data and model lifecycle and across the data ecosystem, which may include researchers, data scientists, model developers, end users, care providers, and patients, among others. Participants then considered how transparency could serve each stakeholder's information and decision needs.

Transparency itself is a means to an end, and this led participants to consider other crucial issues such as quantifying the fitness of a dataset or a model for use and reuse, detecting and mitigating bias in a dataset or a model to prevent unlawful discrimination and improve health equity, and enacting the principles of data protection to ensure that patients can control how their data are used, while, at the same time, benefit from cutting-edge research.

The goals of this workshop were to (1) assess the current landscape of responsible AI in NIH-sponsored research; (2) identify critical capability gaps – in terms of datasets, methods, technical tooling, organizational processes, ecosystem incentives, and workforce readiness – that impede responsible AI adoption through the ecosystems; and (3) lay the groundwork for developing transparency guidelines for awardees who use, curate or produce AI-ready datasets, pre-trained models, or AI tools and methods in their NIH-funded projects.

Overview and Highlights

The workshop agenda included presentations from ODSS and other experts in the field, plenaries, and breakout sessions, all fostering interactive discussions.

[1] The authors define the term "responsible AI" as incorporating ethics, legal compliance, trustworthiness, and safety into the design, development, and use of AI.

ODSS Director, Dr. Susan Gregurick, welcomed participants and introduced the updated NIH Strategic Plan for Data Science, 2023–2028. She highlighted various efforts of accelerating trustworthy AI, including the 2023 Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. Dr. Laura Biven laid the groundwork for the workshop, discussing the goals and objectives. Invited experts: Dr. Julia Stoyanovich, Institute Associate Professor, Department of Computer Science and Engineering, Tandon School of Engineering, and the Center for Data Science, New York University; Dr. Tina Hernandez-Boussard, Associate Dean of Research and Professor of Medicine (Biomedical Informatics), Biomedical Data Science, Surgery, and (by courtesy) Epidemiology and Population Health, Stanford Medicine, Stanford University; and Dr. Aaron Lee, Associate Professor, C. Dan and Irene Hunter Endowed Professor, Department of Ophthalmology, University of Washington—elaborated on data/model transparency, what could go wrong with AI in biomedical and behavioral research, and the current state of data and model transparency, respectively.

The main component of the event was four working sessions that were thematically structured around case studies. These were undertaken in parallel by five breakout groups, each with a thematic lens that focused on a particular case study. The case studies were identified ahead of time by the organizing team to highlight the importance of the socio-technical context (see Appendix 2):

1. Synthetic Data: This breakout focused on synthetic data—how they are generated, how they might be used, and how they could have both positive and negative impacts on human health. Discussions considered the need for and the challenges related to synthetic data, including fitness for use, bias, ethical concerns, and generalizability.
2. Data Sharing for General Reuse: This breakout discussed the responsible reuse of shared data for AI, which requires technical, operational, ethical, privacy, and regulatory considerations to assess whether the data are fit for purpose.
3. Multimodal Data: This breakout discussed using multimodal data within AI model development, validation, and translation for clinical implementation (e.g., combining structured data, such as diagnoses, with unstructured data, such as text or images).
4. Foundational Models: This breakout explored key concepts behind these models, as well as the implications when creating and using these models in multiple settings involving the care provider, patient, researcher, developer, and community as a whole.
5. Proxy Variables: This breakout examined the use of proxy variables in algorithms. Proxy variables are confounders and, therefore, are used (intentionally or unintentionally) in place of another variable that has a true causal relationship with the outcome.

All five of the breakout groups followed the four sequential working sessions described in the agenda:

- Session 1—*Exploring Use Cases and Stakeholder Mapping Breakouts*—on January 31, 2024, identified relevant stakeholders and their information needs, concerns, and decisions.
- Session 2—*Achieving Transparency and Mapping Capability Gaps*—on February 1, 2024, focused on mapping stakeholder needs to information sources and identifying capability gaps.
- Session 3—*Guidance*—on February 1, 2024, engaged participants in developing draft guidance and best practices for actions taken by NIH awardees.

- Session 4—*Future Trends*—on February 2, 2024, identified trends in AI transparency. Each breakout session was followed by a readout plenary and discussion. The workshop's five breakout themes—synthetic data, data sharing for general reuse, multimodal data, foundational models, and proxy variables—spanned the spectrum of the AI life cycle from data collection to prediction reasoning.

ODSS staff, workshop leads, NIH-wide AI Ethics Working Group members, and other representatives from co-sponsoring NIH institutes and centers moderated the breakout sessions and plenary readouts. More than 150 participants registered and attended the workshop, representing a diverse range of backgrounds, including academia and research institutions, healthcare and hospital organizations, industry and corporate sectors, non-profit organizations, government and public sectors, medicolegal fields, and others. NIH anticipates that this workshop will help establish an ethical framework for the responsible use of AI in biomedical and behavioral research, compile guidance to promote and improve transparency in this research area, and be a catalyst to inform solutions and future studies.

Key Themes: Summary

This workshop took a practical, stakeholder-based approach to transparency. Participants identified stakeholders[2], which may include researchers, data scientists, model developers, end users, patients, and members of the care provider team. Participants then considered how transparency could serve the stakeholders' information and decision needs. The key themes that emerged from the interactive discussions during the workshop are summarized below and are elaborated on in the remainder of this document.

Empowerment, engagement, and education of patients, communities, researchers, and practitioners

Participants recognized that enhancing the effectiveness, safety, and inclusivity of AI-driven research will require strategies that empower patients, communities, researchers, and care providers of all levels to be actively involved in the decision-making processes from the outset[3] of the AI lifecycle. This, in turn, requires the development and broad dissemination of new training, education, and public engagement methods that go hand-in-hand with the imperative of transparency and support the objectives of human agency and control.

[2] The authors define the term "stakeholder" as an individual or a group who is impacted by data collection or use, directly or indirectly.

[3] Chin, Marshall H., et al. "Guiding principles to address the impact of algorithm bias on racial and ethnic disparities in health and health care." JAMA Network Open 6.12 (2023): e2345050-e2345050. https://doi.org/10.1001/jamanetworkopen.2023.45050

The "why" and "how" of transparency

Participants observed that while documentation methods like "datasheets for datasets"[4], "health sheets"[5], "model cards"[6], and "nutritional labels for data and models"[7] are starting to gain popularity, there are currently no documentation standards and hardly any tooling to reduce the burden of manual documentation. Coupled with a lack of institutional and ecosystem-wide incentives, datasets and models are documented inconsistently and only sporadically. This limits the potential for the safe and effective reuse of AI-ready datasets and pre-trained models today and also makes future efforts to develop lifecycle-wide support for data and model documentation challenging. The participants suggested that a metadata registry (similar to clinicaltrials.gov for randomized clinical trials) for datasets and AI models should be constructed to document and record enduring artifacts supporting the full development and iterative lifecycle[8].

Some specific properties of responsible AI – such as data security and privacy (in the sense of de-identification) and auditing models for bias so as to prevent unlawful discrimination – are already receiving attention[9,10]. However, other properties, including transparency, robustness, and safety, remain inadequately supported. Further, there are currently insufficient methods that treat these properties *holistically*, both by considering them as complementary objectives and by treating them throughout the lifecycle of problem identification, data collection and curation, model training, data and model re-use, and post-deployment monitoring, rather than separately, in specific pipeline stages.

From Today's Best Practices to a Future Vision

Substantial progress was made during the workshop towards establishing guidance for NIH awardees, with the recognition that the transparency interventions that awardees can enact today – despite the current limitations in transparency tooling and the lack of clear incentives for transparency – will go a long way toward building an ecosystem of responsible data and model reuse in years to come. Further,

---

[4] Gebru, Timnit, et al. "Datasheets for datasets." Communications of the ACM 64, 12 (Dec 2021), 86-92. https://doi.org/10.1145/3458723

[5] Rostamzadeh, Negar, et al. "Healthsheet: Development of a Transparency Artifact for Health Datasets." arXiv:2202.13028 [cs.AI] (2022). https://doi.org/10.48550/arXiv.2202.13028

[6] Mitchell, Margaret, et al. "Model Cards for Model Reporting." Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19), Association for Computing Machinery, New York, NY, 220-229. https://doi.org/10.1145/3287560.3287596

[7] Stoyanovich, Julia; Howe, Bill. "Nutritional Labels for Data and Models." Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, Vol 42 No 3, 13-23 (Sept 2019) http://sites.computer.org/debull/A19sept/p13.pdf

[8] Ng, Madelena Y., et al. "The AI life cycle: a holistic approach to creating ethical AI for health decisions." Nat Med 28.11 (2022): 2247-2249. https://doi.org/10.1038/s41591-022-01993-y

[9] Dignum, Virginia. "Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way." Springer Cham (2019). https://doi.org/10.1007/978-3-030-30371-6

[10] Siala, Haytham; Wang, Yichuan. "SHIFTing artificial intelligence to be responsible in healthcare: A systematic review." Social Science & Medicine, Vol 296, 2022, 114782, ISSN 0277-9536. https://doi.org/10.1016/j.socscimed.2022.114782

it was observed that substantial progress towards responsible AI in biomedical and behavioral research can only be achieved with the help of a comprehensive data and AI governance framework.

Theme 1: Empowerment, engagement, and education of patients, communities, researchers, and practitioners

Empowering patients and their communities

Transparency is an enabler of patient and community engagement. Enhancing the effectiveness, safety, and inclusivity of AI-driven research will require strategies that empower patients, participants, and communities[11] to be actively involved in the decision-making processes from the outset. Mechanisms that seek to integrate their perspectives, needs, and insights in AI research governance will improve the alignment of future research outputs and drive transparency requirements as well as broader ethical decisions and considerations. Developing and adhering to these participatory AI research governance structures will improve the likelihood that ethical considerations and the diverse needs of all stakeholders are identified and integrated throughout the AI lifecycle. This model of engagement will enhance the range of insights for AI research while also improving trust and accountability between AI developers, researchers, and the communities they aim to serve, ultimately leading to more equitable, effective, and sustainable AI solutions.

Opportunities and challenges

Workshop participants highlighted the need for *creating more direct opportunities for engagement and input* from key stakeholders, particularly patients and their communities, centered on co-design, to understand better the benefits, costs, and societal impact of the use of AI models in healthcare settings. Doing so would involve seeking input from the data subjects – a category of stakeholders from whom data is being collected — on issues surrounding *data collection, use, and reuse*, with a particular focus on benefit sharing when data is used and reused and on data sovereignty. This input should be sought and documented for transparency *across the healthcare ecosystem* and *continuously throughout the lifecycle* of data collection and use and model development and deployment to ensure that the use of AI is meeting the needs of the patients and their communities.

To improve patient engagement and input related to ethical and transparent AI in healthcare, participants identified a need to *improve education* to all communities about the opportunities and challenges related to the development and use of AI, as well as an understanding of how their data – generated through hypothesis-driven research studies or through other means like routine healthcare – could be used in AI. This includes educating patients and communities about how the consent process may impact the use of their data in AI-related research.

[11] Bondi, Elizabeth, et al. "Envisioning communities: a participatory approach towards AI for social good." Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. 2021. https://doi.org/10.48550/arXiv.2105.01774

Workshop discussions alluded to a long-term vision for robust participatory development and assessment practices for data and AI models, whereby stakeholder needs and preferences influence data collection and model design and help contextualize transparency. Participants also noted that a variety of methods of participatory engagement should be studied to better understand their impact on transparency and on the responsible and ethical development and use of AI.

Training current and future researchers and practitioners

Researchers and practitioners play an essential role in the responsible design, development, use, and oversight of AI in biomedical and behavioral fields. With the increasing adoption of AI in research and clinical practice, and with the increasing sophistication of AI-based tools, there is a pressing need to train current and future researchers and practitioners regarding the capabilities and the limitations of this technology - both in general and specifically in their domain of engagement[12]. This training is necessary because, without an understanding of the capabilities and shortcomings of AI, transparency interventions will be insufficient for accountability, and these issues are only compounded by the rapid evolution of the technology.

Opportunities and challenges

Workshop participants underscored the importance of *studying the evolving nature of the interactions between humans and AI in research and practice*. They recognized the need to understand the landscape of the ever-broader adoption of AI and its impacts on the workforce, and to prepare the workforce for these impacts.

Relatedly, workshop participants recognized the need to *expand education and training* for biomedical researchers and practitioners. This, in turn, necessitates the development of *educational and training programs* geared towards the unique needs of researchers and practitioners. The content of these programs should be customized to the domain of research or practice, as well as the specific role of the trainee. For example, researchers who collect, curate, and share data, models, and other data products (i.e., those who are both *consumers* and *producers*) need to receive training on the importance of transparency as an enabler of responsible data and model sharing and reuse, as well as of other dimensions of responsible AI, such as privacy and data protection, and the quantification and mitigation of under-representation and other types of bias in data and models. These researchers then need to be given access to and knowledge about tools and frameworks that can support them in enacting transparency, interrogating datasets and models for bias and other dimensions of *fitness for use*, and enacting appropriate levels of data protection when publishing their results.

[12] Hernandez-Boussard, Tina, et al. "Promoting Equity In Clinical Decision Making: Dismantling Race-Based Medicine: Commentary examines promoting equity in clinical decision-making." Health Affairs 42.10 (2023): 1369-1373. https://doi.org/10.1377/hlthaff.2023.00545

Workshop participants observed that, while there is a consensus that such education and training is needed, there are very *few materials and methodologies* available for this training. Furthermore, it was noted that there are currently *very few qualified trainers* – individuals with expertise in the application domain, in the technical tools and frameworks, and in responsible AI. Training for ethical AI is challenged by a lack of consensus methodologies. Finally, approaches to mitigation strategies when disparities or other harms resulting from AI vary in effectiveness and often depend on context[13,14]. Participants noted that a compendium[15] of known ethical challenges and pitfalls could be instructive, especially if it were a living registry with best practices and remediations.

Building ecosystem capacity in responsible AI

Improving the effectiveness, safety, and inclusivity of AI-driven research and practice requires the establishment of a *distributed accountability* regime in which all stakeholders contribute to the responsible design, development, use, and oversight of technology. To make progress towards meeting this challenge, there is currently an urgent need to build institutional capacity in responsible AI to support specific efforts on education, training, public engagement, standardization, etc., as well as to *coordinate* these efforts across the ecosystem and connect the stakeholders. The latter function is particularly important because, according to the workshop participants, the goals of effectiveness, safety, and inclusivity cannot be achieved without interdisciplinary collaboration and stakeholder dialog.

Opportunities and challenges

Workshop participants recognized the pressing need to *build the capacity of Institutional Review Boards (IRBs)* and extend the types of reviews they conduct to include ethical, technical, and transparency considerations specific to AI. They looked to the NIH to provide guidance to IRB committees in light of the ever-broader adoption of AI and the widespread collection, sharing, and re-use of AI-ready datasets.

Further, participants underscored the need to *develop risk review and data protection plans* at the grant proposal stage, as well as safety and feedback mechanisms for active monitoring of unintended consequences to close the loop after the deployment or implementation of AI models.

Finally, and related to the opportunities and trends for training current and future researchers, workshop participants articulated the need to *develop standard resources and training* for biomedical and health researchers, similar to CITI training, centered on the societal, ethical, and transparency requirements for datasets and AI models used in their research.

---

[13] Siddique, Shazia M, et al. "The Impact of Health Care Algorithms on Racial and Ethinic Disparities: A Systematic Review." Ann Intern Med. 2024;177:484-496. [Epub 12 March 2024]. https://doi.org/10.7326/M23-2960

[14] Obermeyer, Ziad, et al. "Algorithmic Bias Playbook." Center for Applied AI at Chicago Booth. June 2021. https://www.chicagobooth.edu/-/media/project/chicago-booth/centers/caai/docs/algorithmic-bias-playbook-june-2021

[15] Greene, Kristen K, et al. "Avoiding Past Mistakes in Unethical Human Subjects Research: Moving from Artificial Intelligence Principles to Practice." Computer, vol 57, no 2, pp 53-63, Feb 2024. https://doi.org/10.1109/MC.2023.3327653

Future trends

Participants underscored the importance of enhancing the full lifecycle of AI development through collaborative approaches to human factors, implementation science, and co-design principles by bringing together and training cross-disciplinary expertise. This will require a coordinated effort in public engagement, practitioner and researcher training, and building institutional capacity.

In the spirit of case studies-based deliberation, participants discussed several potential use cases of AI that may result in substantial benefits but that also require new and more robust methods for engaging and empowering stakeholders – including but not limited to informed consent – to identify and control potential risks. These included "AI counselors" – individuals who can interpret AI results within the context of the clinical, explain their limitations, and inform participants about data sharing AI development.

Theme 2: The "why" and the "how" of transparency: incentives, standards, and tooling

Our collective ability to realize the full potential of transparency to accelerate discovery and improve patient experience and outcomes hinges – in large part – on the incentives for transparency. For example, why would an individual data or model creator engage in the difficult work of creating datasheets or model cards and keeping these documentation artifacts up to date as their datasets or models evolve? These incentives can come in the form of direct benefits of transparency to one's research and clinical practices, in the form of institutional and community recognition, or in the form of legal, regulatory, or other types of mandatory requirements. The effectiveness of these incentives largely depends on the available standards and tools, as well as the processes in which transparency efforts are integrated.

Further, participants observed that while documentation methods like datasheets for datasets, health sheets, model cards, and nutritional labels for data and models are gaining popularity[16], there are currently no documentation standards and hardly any tooling to generate and maintain documentation semi-automatically. Coupled with a lack of institutional and ecosystem-wide incentives, datasets and models are documented only sporadically. This limits the potential for the safe and effective reuse of AI-ready datasets and pre-trained models today and also makes any future efforts to develop lifecycle-wide support for data and model documentation challenging.

Identifying barriers and providing incentives for transparency

Workshop participants reported that transparency is becoming increasingly important with the ever-broader adoption of AI models and their ever-increasing complexity. Without transparency, AI models

[16] Castaño, Joel, et al. "Analyzing the Evolution and Maintenance of ML Models on Hugging Face." IEEE/ACM 21st International Conference on Mining Software Repositories (MSR). IEEE, 2024. arXiv:2311.13380 [cs.SE] (2024). https://doi.org/10.48550/arXiv.2311.13380

will not be trusted by frontline care providers or by patients and their communities. In other words, the most essential incentive for transparency is that, without it being implemented effectively, AI simply cannot be integrated into real-world decision-making.

Further, workshop participants underscored that using AI across scientific processes could introduce limitations and biases, including in clinical trial enrollments and systematic reviews. They saw transparency interventions as helpful for surfacing actual or potential biases in datasets and models.

Opportunities and challenges

In light of the importance of transparency, workshop participants recognized the need to study the current data and model reuse practices to identify organizational, cultural, structural, and technical barriers to transparency. Gaining a better understanding of the current practices and of *the intricate trade-offs between transparency and other aspects of responsible data and model reuse* like privacy and fairness – can help ground the necessary work of building a structure of incentives ("carrots") and requirements ("sticks") to enhance transparency.

Understanding the *trade-offs between transparency on the one hand, and privacy on the other hand*, was seen as a crucial need, coupled with an evaluation of the potential of using *privacy-preserving synthetic data* as an enabler of transparency, to overcome privacy concerns as a barrier to transparency.

Security risks, the "hacking" of medical AI systems, or deliberate misuse of data and models were all seen as additional substantial barriers to transparency. Workshop participants underscored the need to understand the actual security risks that may arise because of transparency interventions and to develop approaches for controlling these risks. A specific issue raised by the workshop participants related to the difficulty – or even the inability – to discern synthetic data from real data, raising concerns over the potential of deliberate misuse of "deepfake" data to mislead or cause other types of harm.

The ever-broader use of *synthetic datasets* was extensively discussed at the workshop, and several additional ecosystem needs surfaced as a result of that discussion. Participants underscored that further research into the ownership, intellectual property, and privacy and security concerns with synthetic data creation and usage is needed to support the responsible use of this data and inform the development of transparency standards and tooling.

While these challenges and uncertainties are being addressed, the workshop participants discussed the value of instituting incentives for greater transparency now, as this will provide the community with useful experience and allow for the iterative improvement of transparency practices. For example, making datasheets, model cards, and nutrition cards requirements for the submission of data or models to repositories or as requirements for publication in peer-reviewed journals.

Metadata Registry for data, model, and lifecycle transparency

Workshop participants discussed the importance of tools and standards for transparency, articulating the need to evaluate the usefulness of existing tools and standards for biomedical and behavioral research, extending the tools and standards to be responsive to the unique needs of this domain, and developing lifecycle-wide support for more automated transparency interventions. Further, workshop participants recognized that transparency could be a powerful enabler of responsible AI, broadly construed, and that it must be considered within the broader set of objectives, such as privacy and security and fairness and equity, with an understanding of the trade-offs between these objectives.

Opportunities and challenges

Workshop participants identified several broad categories of ecosystem needs, namely, catalogs and registries[17]; evaluation metrics and methods; transparency standards; and methodologies and tools to generate transparency meta-data.

Workshop participants highlighted the need to create and maintain *an enduring metadata registry for datasets, models, and other data products*, in which all resources are associated with a DOI and with comprehensive transparency documentation[18]. The development of such a registry substantially challenges the state of the art in meta-data standards and tooling. It requires the refinement of existing metadata approaches, such as datasheets[19], model cards[20], and nutritional labels[21,22], to incorporate domain-specific information[23]. Further, it requires the development of novel lifecycle-centric provenance generation and tracking methodologies that semi-automatically compute and propagate metadata across lifecycle stages and generate a holistic "nutritional label" for a dataset or a model[24]. For example, transparency methods will need to support the recording and tracking of *participant consent*,

[17] A registry is a collection of information about individuals, usually focused around a specific diagnosis or condition.

[18] Chung, Caroline; Jaffray, David A. "Cancer Needs a Robust 'Metadata Supply Chain' to Realize the Promise of Artificial Intelligence." Cancer Res (2021) 81 (23): 5810-5812. https://doi.org/10.1158/0008-5472.CAN-21-1929

[19] Gebru, Timnit, et al. "Datasheets for datasets." Communications of the ACM 64, 12 (Dec 2021), 86-92. https://doi.org/10.1145/3458723

[20] Mitchell, Margaret, et al. "Model Cards for Model Reporting." Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19), Association for Computing Machinery, New York, NY, 220-229. https://doi.org/10.1145/3287560.3287596

[21] Stoyanovich, Julia; Howe, Bill. "Nutritional Labels for Data and Models." Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, Vol 42 No 3, 13-23 (Sept 2019) http://sites.computer.org/debull/A19sept/p13.pdf

[22] Chmielinski, Kasia S, et al. "The Dataset Nutrition Label (2nd Gen): Leveraging Context to Mitigate Harms in Artificial Intelligence." arXiv:2201.03954 [cs.LG] (2022). https://doi.org/10.48550/arXiv.2201.03954

[23] Rostamzadeh, Negar, et al. "Healthsheet: Development of a Transparency Artifact for Health Datasets." arXiv:2202.13028 [cs.AI] (2022). https://doi.org/10.48550/arXiv.2202.13028

[24] Stoyanovich, Julia; Howe, Bill. "Nutritional Labels for Data and Models." Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, Vol 42 No 3, 13-23 (Sept 2019) http://sites.computer.org/debull/A19sept/p13.pdf

help reason about data ownership, and substantiate its responsible and legally compliant use, sharing, and reuse.

Provenance, defined as tracking the data to its origin, should record and manage methodologies throughout the whole data cycle and be implemented with the help of controlled vocabularies, standards, and software tools and platforms. Data provenance should be independently verifiable, with a mechanism set up for third-party verification. An additional requirement for the metadata registry was to implement *transparency on the computational and environmental costs of training and deploying AI models*.

Relatedly, workshop participants articulated the need to develop *guidelines and standards for data and metadata quality management* as part of the metadata registry. These quality management standards would be based on feedback from multidisciplinary stakeholders, including ethicists, data scientists, clinicians, and patients or their representatives. They would include change control management to remain current after a model is deployed, explicitly linking to any information about proxy variable use, and bias mitigation to prevent related harms.

Further, workshop participants articulated the need to develop methodologies and tools to *assess the fitness for the use of a dataset or a model for a specific task*. They specifically discussed the need to develop such tools and methodologies for synthetic datasets and models trained on such datasets. Such assessment will be based on provenance information accompanying the synthetic dataset and clearly and explicitly document its origin and history of derivation, as well as the primary purpose for which it was produced. Notably, as underscored by workshop participants, assessment of fitness for use cannot be done retroactively (i.e., without any provenance information and based on the synthetic data alone), and it cannot be done in a task-agnostic manner.

Metadata guidelines and standards, data and metadata quality management, and quantification of fitness for use were all seen as synergistic with the need to *develop and standardize evaluation metrics and methods* for datasets, models, and other data products. Results of this evaluation would be surfaced through transparency labels, and they would be informed by the properties of the datasets (e.g., how a dataset was collected or synthesized, whether proxies were used) and of the models (e.g., what were the optimization objectives during model training), as well as by the anticipated context of use. Workshop participants stated that the development and standardization of evaluation metrics and methodologies should include assessment strategies of black box models, uncertainty quantification, detection and mitigation strategies for proxy variables used, and metrics for synthetic data usage.

Participants frequently spoke about the need to develop and maintain a range of references for the community, which would be easily accessible and integrated with the meta-data registry. Importantly, these references must come with robust community curation mechanisms, akin to Wikipedia, with incentives to ensure that the information they contain is factually correct (or that it represents community consensus) and that it is based on input from a diversity of stakeholders.

One type of resource that would become an integral part of the metadata registry is a *living registry of proxy variables*. A proxy variable is a variable that is used in place of another variable to draw inferences or make recommendations, either unintentionally or because that other variable is unavailable in the dataset or difficult to measure. A commonly utilized example is race, which is often a proxy for other variables, such as social determinants of health, genetic factors, or immigration status[25]. Another well-described example was the use of recent healthcare costs as a proxy for healthcare needs[26]. The use of proxy variables, while at times unavoidable, has been shown to reinforce pre-existing bias[27]. To ensure the responsible and equitable use of data and AI in biomedical and behavioral research and practice, it is crucial to document the context of data collection and identify known uses of proxy variables in a registry. This registry should further contain information about the known biases encoded in the proxy variables due to the context of data collection or to the context of use, and it should document instances of harm and bias mitigation strategies.

Another needed resource is *an "incidents database,"* cataloging examples of harms that resulted from the use of datasets or models in biomedical and clinical research and practice. This resource can be modeled after the AI Incident Database[28], but it must include *substantial additional human review and fact-checking* to ensure that the information it contains is factually correct.

Future trends

In addition to the immediate ecosystem needs, participants underscored the new challenges inherent in the impending adoption of generative AI technologies (such as those based on large language models, or LLMs) into clinical and biomedical research and practice. The assessment of these technologies is particularly difficult for several reasons.

First, the relationship between the properties of the training corpus, on the one hand, and the correctness and robustness of model outputs, on the other hand, is not currently well understood. Further, it is not clear whether and how to incorporate rule-based reasoning and, more generally, domain context into the operation of these models, how to measure and mitigate bias and uncertainty, and how to assess and mitigate privacy and security risks due to the use of these models. Taken together, these challenges put into question our current ability to put generative AI into safe use in biomedical and behavioral research and practice.

Addressing these issues requires a stakeholder-centered approach in which evaluation metrics and standards are developed with community input and where technical mitigations follow these community-generated requirements. In the context of generative AI and beyond, workshop participants

[25] Siddique, Shazia M, et al. "The Impact of Health Care Algorithms on Racial and Ethinic Disparities: A Systematic Review." Ann Intern Med. 2024;177:484-496. [Epub 12 March 2024]. https://doi.org/10.7326/M23-2960
[26] Obermeyer, Ziad, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." Science 366, 447-453 (2019). https://doi.org/10.1126/science.aax2342
[27] Friedman, Batya; Nissenbaum, Helen. "Bias in computer systems." ACM Trans. Inf. Syst. 14, 3 (July 1996), 330-347. https://doi.org/10.1145/230538.230561
[28] https://incidentdatabase.ai/

underscored the importance of making stronger connections between data scientists and model developers and a new society of experts in other scientific disciplines (including sociology, ethics, and anthropology) who could serve as collaborators for AI developers.

The need for new tools was also discussed. For example, new tools and methods for characterizing uncertainty in generative AI applications. For example, generative AI can be used to augment or rebalance a cohort of data and better tools are needed to capture the impact of using those data for training AI models. Another example is where generative AI is used to improve the fidelity of medical images and uncertainty at the pixel level can inform the interpretation of those images and the resulting diagnoses.

Participants also discussed the need to characterize black-box foundation models better. Performance metrics, context, and provenance are critical in these cases for responsible and ethical reuse.

Workshop participants also saw the need to develop new socio-technical methodologies for the detection of both evident latent inequalities in healthcare delivery and for reparative model development to compensate for these inequalities.

Concluding remarks and future vision: Towards an ecosystem of responsible AI

A unified approach to AI research guidance is needed across the diverse scientific sectors involved in AI research and development, including academic communities, industry, health systems, and federal funding and regulatory agencies such as the National Institutes of Health (NIH), National Science Foundation (NSF), Department of Energy (DOE), and National Institute of Standards and Technology (NIST), among others. Because outputs from one domain of AI research are often rapidly integrated into other domains, this harmonization should extend beyond the initial Request for Information (RFI) to encompass the full lifecycle of AI development across scientific domains. Ensuring adequate alignment across these diverse research entities and agencies will foster a more cohesive and efficient framework for AI innovation that improves collaboration and the development of robust tools. Without harmonized guidance across these communities, the development of effective, safe, and ethical AI will be hampered and less generalizable across domains.

Based on the input from workshop participants, the *following recommendations have emerged for NIH awardees*. The current ecosystem of data and AI use is highly heterogeneous and complex in terms of stakeholders and their goals, research questions and clinical needs, and resources, standards, and tooling. Because of this heterogeneity and complexity, these recommendations are for actions that are feasible to undertake *today*, given the current state of the art in transparency tools and standards. Importantly, following these recommendations will help the community move towards an ecosystem of responsible, equitable, and ethical use of AI in biomedical and clinical research and practice, accelerating

our collective ability to conduct research and move its results from bench to bedside, all with the goal of improving patient outcomes.

During the workshop, it was observed that substantial progress towards responsible AI in biomedical and behavioral research can only be achieved with the help of a comprehensive data and AI governance framework. This framework should be sufficiently robust to accommodate the multitude of use cases, responsible AI objectives, and stakeholder needs while being sufficiently flexible and lightweight not to slow down research and innovation. It should be grounded in the accepted ethical principles that underlie biomedical and behavioral research and complemented by technical innovation, workforce development, and patient and community education to lower the friction of broad adoption.

Appendix 1: Guidance for Awardees

Data and Model Creation/Collection

- **Cross-disciplinary, multi-stakeholder teams**. Include individuals with expertise in bioethics in the NIH grant personnel list and enhance collaboration with sociologists, historical experts, anthropologists, ethicists, experts in bias/fairness, and philosophers as part of both key personnel and advisory boards. Establish practices and engagements with patients, communities, advocates, frontline care workers, and other relevant stakeholders. Such collaborations are essential for identifying potential individual and social harms and appropriate risk-based strategies.
- **Inclusive Approach to Data**. Collect data on *self-reported* race and ethnicity, including granular ethnicity data categories. Also, collect additional data so they can be included in future models, including but not limited to:
    - Country of origin
    - Social factors (insurance status, zip code, education, parental educational status and income, credit history, social media habits, home ownership, licensures, internet access, car ownership, rent burden (% of income going to rent)) with an understanding that these may be sensitive issues
    - Everyday Discrimination Scale to measure effects of systemic racism
    - Language
- **Use datasheets or similar documentation** and flag proxy variables known to cause harm, state the purpose of data collection and intended use/application, as well as ethical considerations.

Data Sharing

- **Use Persistent, Unique Identifiers (PUIDs) for shared data.** PUIDs, such as DOIs, enable accurate referencing, versioning, and linking among data, models, and publications.
- **Publish checksums for shared data.** Checksums provide a first level check on information integrity for downloads.
- **Data Dictionaries and Proxy Variables.** Include data dictionaries for variables and data types, calling out variables that are used as proxies for other information.
- **Create and share datasheets or similar documentation.** At a minimum, this documentation should describe the motivation, composition, collection process and consent, pre-processing, anticipated use cases, known assumptions, limitations for reuse, risk and bias assessments, and other information relevant for ethical reuse.
    - Include an **explicit description of the conceptual model** for the underlying biological or health system and the extent to which the shared data capture, either directly or through proxy variables, behaviors of interest.
    - Disclose and document the use of LLMs or other **generative AI capabilities** when used.
    - Include **metrics** that describe relevant biases, imbalances, sparseness, or uncertainties in the data.

    - For **synthetically created datasets**, include as much detail as possible, including the anticipated best- and worst-case scenarios, uncertainties, descriptives of distributions of data generated with normative ranges, whether counterfactuals were considered, the mechanism of synthesis, what scientific constraints were taken into account, and what forms of human engagement and oversight were performed.
- **Update Information**. Update metadata artifacts if new gaps, limitations, or flaws are discovered.

Model Training

- Ensure models are **hypothesis-driven** with a statement of ethical/responsible AI intent; for example, report the role of each variable in the model a priori. Reparative models should be strongly encouraged.
- **Avoid using race and ethnicity in prediction models** unless it is justified by an explanation of what race and ethnicity are believed to be proxies for and the availability of more precise variables in the dataset.
- Attempt to **identify potential proxy variables** by determining how input variables vary across racial and ethnic groups prior to model development because it is important to avoid unintended harms of proxy variable utilization. Include patients or patient groups in model creation.
- If bias is discovered, **report mitigation plan** in progress reports to NIH, and Limitations section in NIH-funded manuscripts.

Model Sharing

- **Use Persistent, Unique Identifiers (PUIDs).** PUIDs, such as DOIs, enable accurate referencing, versioning, and linking among data, models, and publications.
- **Publish checksums for shared data.** Checksums provide a first-level check on information integrity for downloads.
- **Create and share model cards or similar documentation.** At a minimum, this documentation should describe the motivation, composition, collection process and consent, pre-processing, anticipated use cases, known assumptions, limitations for reuse, risk and bias assessments, and other information relevant for ethical reuse.
    - Documentation should **include references and PUIDs for underlying training and testing datasets**, as well as standardized testing and processing steps to transform the data. When sharing a tuned or otherwise modified model, include references and PUIDs to the underlying model.
    - Include an **explicit description of the conceptual model** for the underlying biological or health system and the extent to which the shared model could be generalized to other systems.
    - Disclose and document the use of LLMs or other **generative AI capabilities** when used.
    - Include **metrics** that describe model performance, biases, or uncertainties.
- **Describe the compute complexity for training and inference**. Ethical AI needs to account for energy consumption. As a proxy, shared models should be accompanied by estimates for the compute needed for training and inference.

- **Create patient-facing health sheets, data cards,** and model cards for education and consent. The model card should be updated dynamically and verified by a third-party verifier.
- **Update information**. Update metadata artifacts if new gaps, limitations, or flaws are discovered. If a model is later discovered to have an unacceptable performance in a critical area, provide a replacement model and alert downstream users who have used the flawed model.

Data Reuse

- **Data Protection Plan**. Multiple stakeholders, including research funders and data subjects, have an interest and/or responsibility to ensure data management practices appropriately protect subjects' privacy, align with subjects' consent, and avoid social harm. Reusers of data should develop a Data Protection Plan for the entire data and AI lifecycle, addressing data protection, information security, auditing, and governance.
    - The plan should include a **"fit for purpose" evaluation** of the data and potential impacts of imbalances or biases in the dataset.
    - The Data Protection Plan should be **incorporated into proposal Data Management and Sharing Plans (DMSPs)** and documentation accompanying shared data and models.
- **Multidisciplinary approach.** Include individuals with expertise in bioethics and enhance collaboration with sociologists, historical experts, anthropologists, ethicists, bias/fairness experts, and philosophers as part of both key personnel and advisory boards. Such collaborations are essential for reflecting the perspectives of multiple stakeholders and ensuring that data is being reused in an ethical and transparent manner.
- **Chain of Provenance**. Log the reuse of data by re-sharing a new, modified dataset or data product (following the guidance for *Data Sharing*) or by amending the original data documentation.

Model Reuse

- **Model Protection Plan**. Multiple stakeholders, including research funders and data subjects, have an interest and/or responsibility to ensure model management practices appropriately protect subjects' privacy, align with subjects' consent, and avoid social harm. Reusers of models should develop a Model Protection Plan for the entire data and AI life cycle addressing data protection, information security, auditing, and governance.
    - The plan should include a **"fit for purpose" evaluation** of the model and potential impacts of biases.
    - The Model Protection Plan should **be incorporated into proposal DMSPs** and documentation accompanying shared models.
- **Multidisciplinary approach.** Include individuals with expertise in bioethics and enhance collaboration with sociologists, historical experts, anthropologists, ethicists, bias/fairness experts, and philosophers as part of both key personnel and advisory boards. Such collaborations are essential for reflecting the perspectives of multiple stakeholders and ensuring that data is being reused in an ethical and transparent manner.
- **Share Performance and Ethics Findings**. New findings related to the model performance, biases, or information relevant to the informed and ethical reuse of the model should be shared, for

example, through edits, amendments, or comments on the associated model card or other documentation.

- **Chain of Provenance**. Log the reuse of models by re-sharing a new, modified model (following the guidance for *Model Sharing*) or by amending the original model documentation. Ensure any new data used for training or testing are referenced using PUIDs.
- **Information and Data Quality**. Ensure that the transfer of model weights are verified using checksums after download to ensure no bit corruption during transfer.

Model Deployment / Implementation / Application and Continuous Assessment

- **Testing and Assessment for Ethical, Social, Legal Implications**. Ethical deployment of AI requires awareness of model behavior and performance. Testing and assessment of model performance should include assessment of potential biases or imbalances with respect to race, ethnicity, socio-economic status, disease status, and other potentially disadvantaged groups or combinations of groups. Testing and assessment should go beyond data-based assessments to also consider AI as an intervention in complex social structures and assess the ways in which introducing an AI capability can change human behavior.
- **Continual Testing and Assessment**. AI model performance may "drift" as the application circumstances evolve. Similarly, AI models may be continually updated and refined with the incorporation of new data. Risk management requires continual assessments.

Appendix 2: Case Study Descriptions

The workshop organizing committee developed the following case studies to focus the parallel breakout sessions and ensure a breadth of topics were covered. These case studies or themes were used as lenses through which the opportunities and challenges of AI transparency were viewed.

1. **Proxy Variables**
   In this breakout session, we will examine the use of proxy variables in algorithms. Proxy variables are confounders and, therefore, are used (intentionally or unintentionally) in place of another variable that has a true causal relationship with the outcome. A notable example is the use of race and ethnicity in prediction models, as many experts believe that these variables are often oversimplified proxies for such variables as genetic ancestry or complex environmental and social factors. Other examples include the use of healthcare costs as a proxy for healthcare needs (Obermeyer 2019); given that less money is spent on Black patients who have the same level of need as White patients, the examined algorithm falsely concluded that Black patients are healthier than equally sick White patients.
2. **Synthetic Data**
   This breakout focuses on synthetic data—how they are generated, how they might be used, and how they could have both positive and negative impacts on human health. We will discuss specific considerations for the need for and challenges related to synthetic data, including realism, bias, degradation, ethical concerns, and generalizability. Where possible, specific examples will be discussed and used in developing best practices and guiding principles for the ethical use and transparency of synthetic data.
   Additional Resources:
   - Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and External Controls in Clinical Trials - A Primer for Researchers. Clin Epidemiol. 2020 May 8;12:457-467. doi: 10.2147/CLEP.S242097. PMID: 32440224; PMCID: PMC7218288.
   - www.statnews.com/2019/02/05/synthetic-control-arms-clinical-trials/
   - www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-fdas-new-strategic-framework-advance-use-real-world
   - nap.nationalacademies.org/catalog/26922/opportunities-and-challenges-for-digital-twins-in-biomedical-research-proceedings
   - fastmri.org
3. **Multimodal Data**
   This breakout will discuss using multimodal data within artificial intelligence (AI) model development, validation, and translation for clinical implementation (e.g., combining structured data, such as diagnoses, with unstructured data, such as text or images). This will include specific considerations for the need for and challenges of generating and linking data in relation to ethics, bias, privacy, and transparency when combining complex, multimodal data with such details as time-course relevance.

Additional Resources:

- HeLM: arxiv.org/pdf/2307.09018.pdf
- Multimodal wearables: Dunn, J., Kidzinski, L., Runge, R. et al. Wearable sensors enable personalized predictions of clinical laboratory measurements. Nat Med 27, 1105–1112 (2021). doi.org/10.1038/s41591-021-01339-0

4. **Foundational Models**

Foundational models and the closely related Large Language Models (LLMs), such as ChatGPT, have sparked a huge wave of innovations combining the use of AI with the amazing capabilities of these models to integrate and deliver information. This breakout series will explore key concepts behind these models, as well as the implications when creating and using these models in multiple settings involving the clinician, patient, researcher, developer, and community as a whole. Anticipated areas of discussion include such topics as transparency, ethics, privacy, ownership, and reliability.

Additional Resources:

- www.projectpro.io/article/foundational-models-vs-large-language-models/893
- aws.amazon.com/what-is/foundation-models/
- pubmed.ncbi.nlm.nih.gov/37045921/
- pubmed.ncbi.nlm.nih.gov/36523642/
- hai.stanford.edu/news/shaky-foundations-foundation-models-healthcare

5. **Data Sharing for General Reuse**

Data sharing has great potential to accelerate scientific innovation; however, it occurs without knowledge of how, whether, or by whom the data will be reused. Responsible reuse of shared data for AI requires technical, operational, ethical, privacy, and regulatory considerations to assess whether the data are fit for purpose.

Appendix 3: Agenda

Day 1: Wednesday, January 31, 2024

1:00 p.m. – 1:05 p.m.	**Welcome and Logistics** Laura Biven, Ph.D. *Data Science Technical Lead, Office of Data Science Strategy (ODSS), National Institutes of Health (NIH)*
1:05 p.m. – 1:15 p.m.	**Welcoming Remarks** Susan K. Gregurick, Ph.D. *Director of the Office of Data Science Strategy, Associate Director for Data Science, ODSS, NIH*
1:15 p.m. – 1:35 p.m.	**Workshop Goals and Expectations; Stakeholder Mapping/Setting Expectations** Laura Biven, Ph.D. *Data Science Technical Lead, ODSS, NIH*
1:35 p.m. – 1:55 p.m.	**Introduction to Transparency** Julia Stoyanovich, Ph.D. Associate Professor, Department of Computer Science and Engineering, Tandon School of Engineering, and the Center for Data Science, New York University
1:55 p.m. – 2:15 p.m.	**What Could Go Wrong?** Tina Hernandez-Boussard, Ph.D. *Professor of Medicine, Biomedical Data Science, of Surgery and, by courtesy, of Epidemiology and Population Health, Stanford University School Medicine*
2:15 p.m. – 2:30 p.m.	***Break***

2:30 p.m. – 3:50 p.m.	**Use Case Breakout Session #1—Exploring Use Cases and Stakeholder Mapping** ● Synthetic Data: Room 270-A ● Data Sharing for General Reuse: Room 270-B ● Multimodal Data: Room 260-F ● Foundation Models: Room 280-A ● Proxy Variables: Room 150-A
3:50 p.m. – 4:05 p.m.	***Break***
4:05 p.m. – 4:45 p.m.	**Plenary Readout and Discussion**
4:45 p.m. – 5:00 p.m.	**Summary Remarks and Day 1 Closing** Aaron Lee, M.D., M.S.C.I. *Associate Professor, Department of Ophthalmology, University of Washington*

Day 2: Thursday, February 1, 2024

9:00 a.m. – 9:05 a.m.	**Welcome** **Laura Biven, Ph.D.** ***Data Science Technical Lead, ODSS, NIH***
9:05 a.m. – 9:25 a.m.	**The Current State of Data and Model Transparency** Aaron Lee, M.D., M.S.C.I. *Associate Professor, Department of Ophthalmology, University of Washington*
9:25 a.m. – 9:35 a.m.	**Morning Plenary Session—Recap and Expectations** Laura Biven, Ph.D. *Data Science Technical Lead, ODSS, NIH*
9:35 a.m. – 9:50 a.m.	***Break***

9:50 a.m. – 11:15 a.m.	**Use Case Breakout Session #2—Achieving Transparency and Mapping Capability Gaps** ● Synthetic Data: Room 270-A ● Data Sharing for General Reuse: Room 270-B ● Multimodal Data: Room 260-F ● Foundation Models: Room 280-A ● Proxy Variables: Room 150-A
11:15 a.m. – 11:30 a.m.	***Break***
11:30 a.m. – 12:10 p.m.	**Plenary Readout**
12:10 p.m. – 1:15 p.m.	**Lunch Break**
1:15 p.m. – 1:45 p.m.	**Plenary Session by NIH to Provide Guidance for Upcoming Session (NIH speaker TBD)**
1:45 p.m. – 2:00 p.m.	***Break***
2:00 p.m. – 3:15 p.m.	**Use Case Breakout Session #3 Guidance Session** ● Synthetic Data: Room 270-A ● Data Sharing for General Reuse: Room 270-B ● Multimodal Data: Room 260-F ● Foundation Models: Room 280-A ● Proxy Variables: Room 150-A
3:15 p.m. – 3:45 p.m.	***Break***
3:45 p.m. – 4:45 p.m.	**Plenary Readout Session and Discussion** *Hybrid discussion with virtual and in-person participants*

4:45 p.m. – 5:00 p.m.	**Day 2 Closing Remarks** Aaron Lee, M.D., M.S.C.I. *Associate Professor, Department of Ophthalmology, University of Washington*
5:00 p.m. – 5:30 p.m.	**Closed Session for Co-leads and Breakout Chairs.**

Day 3: Friday, February 2, 2024

8:30 a.m. – 8:45 a.m.	**Recap of Day 2** Tina Hernandez-Boussard, Ph.D. *Professor of Medicine, Biomedical Data Science, of Surgery and, by courtesy, of Epidemiology and Population Health, Stanford University School of Medicine*
8:45 a.m. – 9:45 a.m.	**Continued Plenary Discussion on Transparency Guidance and Capability Gaps** *Hybrid discussion with virtual and in-person participants.*
9:45 a.m. – 10:00 a.m.	***Break***
10:00 a.m. – 11:00 a.m.	**Use Case Breakout Session #4—Future Trends (5 breakout rooms)** ● Synthetic Data: Room 150-A ● Data Sharing for General Reuse: Room 270-A ● Multimodal Data: Room 270-B ● Foundation Models: Room 280-A ● Proxy Variables: Room 260-F
11:00 a.m. – 11:15 a.m.	***Break***
11:15 a.m. – 12:00 p.m.	**Plenary Readout Session and Discussion** *Hybrid discussion with virtual and in-person participants.*
12:00 p.m. – 12:15 p.m.	**Closing Plenary Talk and Thanks (NIH speakers and co-chairs)**
12:15 p.m.	**Workshop Adjourns**

Appendix 4: Workshop participants

* Breakout Lead
** Co-chair

Name	Organization
Brian Anderson	MITRE Corporation
Prasanna Balaprakash	Oak Ridge National Laboratory
Vladimir Braverman	Rice University
Thomas Brettin	Argonne National Laboratory
Marino Bruce	University of Houston
* Ansu Chatterjee	U Maryland
Feixiong Cheng	Cleveland Clinic
Marshall Chin	University of Chicago
* Caroline Chung	MD Anderson Cancer Center
Christopher Chute	Johns Hopkins University
Ellen Clayton	Vanderbilt University Medical Center
Issam El Naqa	Moffitt Cancer Center
Nicholas Evans	University of Massachusetts Lowell
Carole Federico	Stanford University
Christopher Gibbons	MD Anderson Cancer Center
** Tina Hernandez-Boussard	Stanford University
Brian Hie	Stanford University
* Maia Hightower	Equality AI
Timothy Hohman	Vanderbilt University Medical Center
Mohammad Hosseini	Northwestern University
Bill Howe	U Washington
* Sajid Hussain	Fisk University
* H V Jagadish	University of Michigan

Joy Jang	University of Michigan, Inter-university Consortium for Political and Social Research (ICPSR)
Xiaoqian Jiang	University of Texas Health Science Center at Houston
William Jordan	American Medical Association
Karuna Joshi	University of Maryland, Baltimore County
* Jayashree Kalpathy-Cramer	University of Colorado
Daniel S. Katz	University of Illinois Urbana-Champaign
Stephanie Kraft	University of Washington
** Aaron Lee	University of Washington
Ashley Lewis	Stanford University
* Vincent Liu	Kaiser Permanente
* Courtney Lyles	University of California, Davis
Bradley Malin	Vanderbilt University Medical Center
Pietro Michelucci	Human Computation Institute
Sean Mooney	University of Washington
Tamra Moore	Prudential Financial
Madelena Ng	Stanford University
Nico Nortje	MD Anderson Cancer Center
Oded Nov	New York University
Lucila Ohno-Machado	Yale University
Shauna Overgaard	Mayo Clinic
Bhavesh Patel	FAIR Data Innovations Hub at the California Medical Innovations Institute (CalMI2)
Desmond Patton	University of Pennsylvania
Jane Pinelis	JHU
Tom Powers	University of Delaware
Arvind Ramanathan	Argonne National Laboratory
Eric Rosenthal	Massachusetts General Hospital

Eugene Santos	Dartmouth College
Katie Shilton	University of Maryland, College Park
* Shazia Siddique	University of Pennsylvania
* Eric Stahlberg	Frederick National Lab
** Julia Stoyanovich	New York University
Tanveer Syeda-Mahmood	IBM Research
Svitlana Volkova	Aptima
Quinn Waeiss	Stanford Center for Biomedical Ethics
* Colin Walsh	Vanderbilt University Medical Center
Gabriella Waters	Center for Equitable AI & Machine Learning Systems - Morgan State University
Wenbin Zhang	Florida International University